

ON²VERS

Online Ontology Versioning



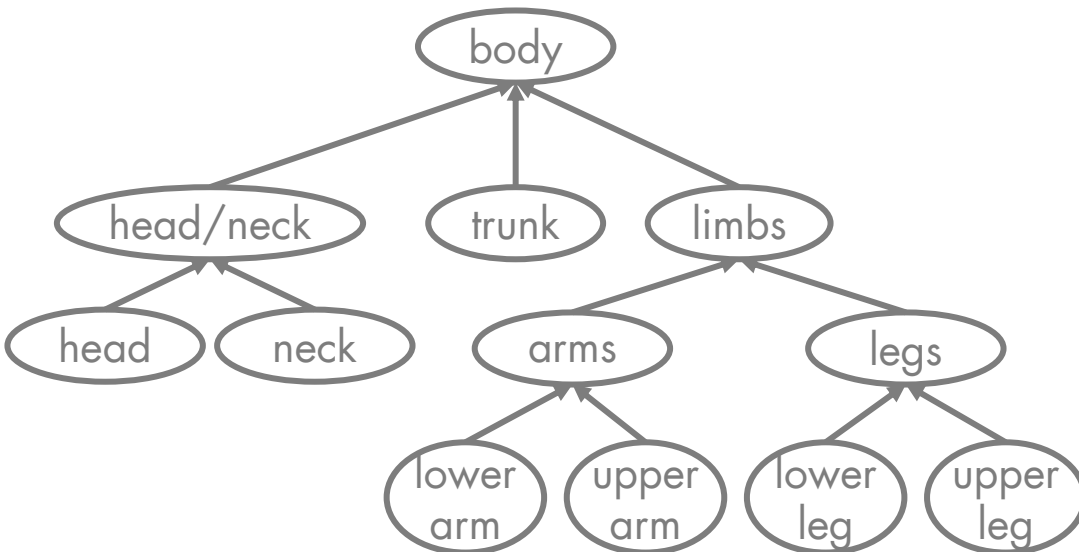
Anika Groß

2nd Network Conference
12th June 2013, Dresden



ONTOLOGIES

- Structured representation of knowledge
- Various applications: ontology-based semantic search, annotations (of "big data"), ...
- Very large ontologies, e.g. in the life sciences



amazon.de Mein Amazon Angebote Gutscheine Verkaufen Hilfe

Alle Kategorien Suche Notebooks

Computer & Zubehör Bestseller Notebooks & PCs Tablets PC-Komponenten PC-Zubehör & M

Kategorie
Computer & Zubehör
Notebooks

Displaygröße
Auswahl aufheben
 28-30 cm (11"-12") (6)
 33-36 cm (13"-14") (71)
 38-41 cm (15"-16") (75)
 43 cm (17") & größer (36)

Computer & Zubehör > Notebooks > Dell > 33-36 cm
1-24 von 71 Ergebnissen

P01308 (INS_HUMAN) ★ Reviewed, UniProtKB/Swiss-Prot
Last modified May 29, 2013. Version 181. History...

Clusters with 100%, 90%, 50% identity | Documents (6) | Third-party data

Names · Attributes · General annotation · Ontologies · Sequence annotation

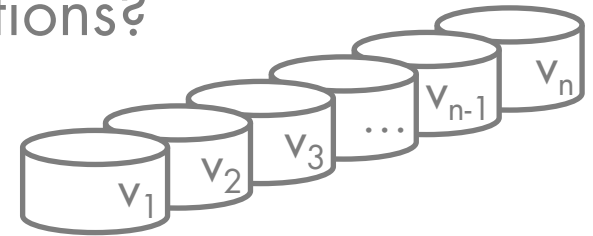
Names and origin

Protein names	Recommended name: Insulin
Cellular_component	Golgi lumen Traceable author statement. Source: Reactome endoplasmic reticulum lumen Traceable author statement. Source: Reactome
Molecular_function	hormone activity Non-traceable author statement (PubMed 1498611) insulin receptor binding Inferred from direct assay (PubMed 7556975). Source: UniProt

Complete GO annotation...

ONTOLOGY EVOLUTION

- Ongoing research, new findings
→ Continuous modifications of ontologies
- Periodical release of new ontology versions
- Impact on dependent data and applications?
(queries, annotation mappings, ontology mappings, ...)



Aims

- Provide services for Online Ontology Versioning (*On²Vers*) in the eScience network
 - Efficient version management, multi-user application
 - Enable basic and complex evolution analyses
 - Migration of ontology-based mappings to current version

PREVIOUS WORK

- Research on evolution of large ontologies
 - GOMMA
 - ContoDiff
 - Codex
 - Evolving ontology regions
- + Efficient version management (internal use only)
- + Basic & complex evolution algorithms/analyses
- Ontology versioning service (multi-user application)
- Mapping migration algorithm
- Application services
- Access via one platform

Kirsten et al. *Journal of Biomedical Semantics* 2011, 2:6
<http://www.jbiomedsem.com/content/2/1/6>



RESEARCH

Open Access

GOMMA: a component-based infrastructure for managing and analyzing life science ontologies and their evolution

Toralf Kirsten^{1,2*}, Anika Gross^{1,2}, Michael Hartung^{1,3} and Erhard Rahm^{1,3}

Journal of Biomedical Informatics 46 (2013) 15–22

Contents lists available at SciVerse ScienceDirect



Journal of Biomedical Informatics

Journal homepage: www.elsevier.com/locate/yjbin



Onto-Diff: generation of complex evolution mappings for life science ontologies¹

Michael Hartung¹, Anika Groß, Erhard Rahm

¹Interdisciplinary Center for Bioinformatics, University of Leipzig, Hitzlerstraße 16-18, 04107 Leipzig, Germany
²Department of Computer Science, University of Leipzig, P.O. Box 100920, 04099 Leipzig, Germany

BIOINFORMATICS APPLICATIONS NOTE Vol. 28 no. 6 2012, pages 856–896
doi:10.1093/bioinformatics/bts029

Databases and ontologies

Advance Access publication January 16, 2012

CODEx: exploration of semantic changes between ontology versions

Michael Hartung^{1,2,*}, Anika Gross^{1,2} and Erhard Rahm^{1,2}

¹Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, Germany and ²Department of Computer Science, University of Leipzig, Leipzig, Germany
Associate Editor: Alair Edarim

Discovering Evolving Regions in Life Science Ontologies

Michael Hartung^{1,2}, Anika Gross^{1,2}, Toralf Kirsten^{1,3}, and Erhard Rahm^{1,2}

¹ Interdisciplinary Centre for Bioinformatics, University of Leipzig

² Department of Computer Science, University of Leipzig

³ Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig
{hartung, tkirsten}@izbi.uni-leipzig.de,
{gross, rahm}@informatik.uni-leipzig.de

Abstract. Ontologies are heavily used in life sciences and evolve continuously to incorporate new or changed insights. Often ontology changes affect only specific parts (regions) of ontologies making it valuable for ontology users and applications to know the heavily changed regions on the one hand and stable regions on the other hand. However, the size and complexity of life science ontologies renders manual approaches to localize changing or stable regions impossible. We therefore propose an approach to automatically discover evolving or stable ontology regions. We evaluate the approach by studying evolving regions in the Gene Ontology and the NCI Thesaurus.

Keywords: ontology evolution, ontology changes, ontology regions.

MAPPING MIGRATION

- Continuous evolution of ontologies requires the adaptation (migration) of their associated mappings
 - E.g., ontology or annotation mappings, ontology-based queries, ...

Scenario for ontology mappings



Semi-Automatic Adaptation of Mappings between Life Science Ontologies

Anika Groß¹, Julio Cesar Dos Reis^{2,3}, Michael Hartung¹, Cédric Pruski², and Erhard Rahm¹

¹ Department of Computer Science, University of Leipzig, Germany

² CR SANTEC, Public Research Centre Henri Tudor, Luxembourg

³ LRI, University of Paris-Sud XI, France

{gross, hartung, rahm}@informatik.uni-leipzig.de,

{julio.dosreis, cedric.pruski}@tudor.lu

Abstract. The continuous evolution of life science ontologies requires the adaptation of their associated mappings. We propose two approaches for tackling this problem in a largely automatic way: (1) a *composition-based adaptation* relying on the principle of mapping composition and (2) a *diff-based adaptation* algorithm individually handling change oper-

- Development of an algorithm for mapping migration

SYSTEM OVERVIEW

Application
Scenarios

Monitoring
Ontology
Changes

Evolution Analyses
for Term Enrichment
Analyses

...



Evolution Analysis
and Mapping
Migration Services

Basic Evolution
Analyses (Statistics,
Search, ...)

Complex Evolution
Analyses (Diff and
Region Detection)

Mapping
Migration



Version
Management
Services


Ontology and
Mapping Versions

Import and Export Services



RDBMS, OBO, OWL, CSV, RDF, XML, ...

ON²VERS - CREATE PROJECTS

- User specific & “open” projects
 - VISO: Visualization Ontology (VISO)
 - Anatomy: Adult Mouse Anatomy Ontology, Fly Anatomy, ...
- Integration of ontology versions into GOMMA repository
 - Import owl/obo files (local, URL)



anikagross Logged in as: anikagross

Source Import <-> Workspace: VISO2

Graphic@VISO

Version: 994, Timestamp: 2012-07-13
Version: 1022, Timestamp: 2012-07-16
Version: 1024, Timestamp: 2012-07-18
Version: 1064, Timestamp: 2012-09-03
Version: 1126, Timestamp: 2012-10-15
Version: 1130, Timestamp: 2012-10-23
Version: 1218, Timestamp: 2013-01-10
Version: 1221, Timestamp: 2013-01-12

Graphic@VISO

[Create New Initial Source](#)

obo owl

File Format

URL Upload

Location Type

How many versions do you like to import?

File for Version 1 Version Timestamp

2013 May						
M	T	W	T	F	S	S
29	30	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2
3	4	5	6	7	8	9

Overview

Projects

Imports

Ontology Import

Statistics

CODEX

BASIC EVOLUTION ANALYSES

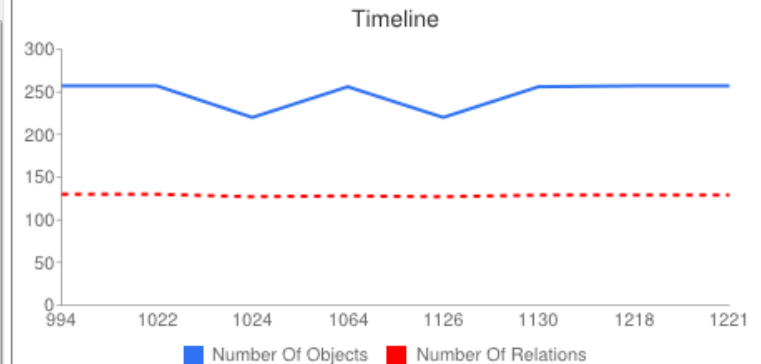


anikagross Logged in as: anikagross

Statistics <-> Workspace: VISO2

Graphic@VISO

Version: 994, Timestamp: 2012-07-13 [Objects: 257, Relations: 130]
Version: 1022, Timestamp: 2012-07-16 [Objects: 257, Relations: 130]
Version: 1024, Timestamp: 2012-07-18 [Objects: 220, Relations: 127]
Version: 1064, Timestamp: 2012-09-03 [Objects: 256, Relations: 128]
Version: 1126, Timestamp: 2012-10-15 [Objects: 220, Relations: 127]
Version: 1130, Timestamp: 2012-10-23 [Objects: 256, Relations: 129]
Version: 1218, Timestamp: 2013-01-10 [Objects: 257, Relations: 129]
Version: 1221, Timestamp: 2013-01-12 [Objects: 257, Relations: 129]



Overview

Projects

Imports

Statistics

Statistics

CODEX

COMPLEX EVOLUTION ANALYSES (CODEX)



Ontology

First version

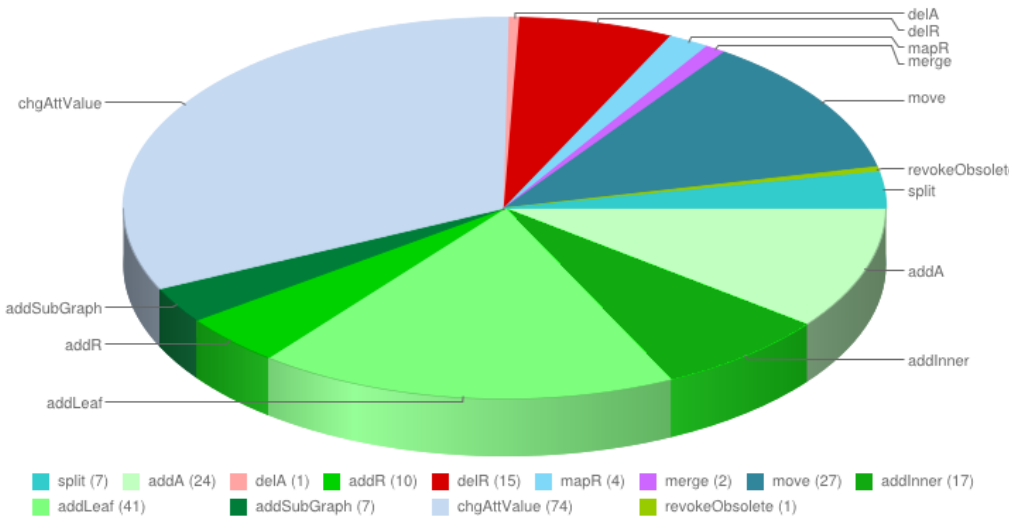
Second version

Change Explorer

addA addInner **addLeaf** addR addSubGraph
chgAttValue definition delA delR is_a mapR merge MOVE
name part_of revokeObsolete split synonym

accessory adipose **amygdaloid** anterior aorta aperture area atrial bladder body **bone** bulb capillary capsule
cerebellum circumventricular collecting **cortex** cortical distal duct epiphyseal epithelium fascia fat
 follicle foot foramen fourth ganglion **gland** greater hand horn hyoid interstitium interventricular
kidney layer left ligament lobe lobule long malleus manubrium medulla middle mouse muscle
 node nucleus olfactory organ outer ovary palmar pararenal perirenal peritubular phalanx plate posterior proximal
 renal septum sphincter tissue **tubule** tunica urinary vein ventricle **vermis** vestibular zone

Compact Diff Result



DONE & ToDo

Application
Scenarios

Monitoring
Ontology
Changes

Evolution Analyses
for Term Enrichment
Analyses

...



Evolution Analysis
and Data
Migration Services

Basic Evolution
Analyses (Statistics,
Search, ...)

Complex Evolution
Analyses (Diff and
Region Detection)

Mapping
Migration



Version
Management
Services


Ontology and
Mapping Versions

Import and Export Services



RDBMS, OBO, OWL, CSV, RDF, XML, ...

FUTURE WORK

- Integrate Mapping Migration Service and Region Explorer into *On²Vers*
- Realize application scenarios
- Integrate user feedback
- Provide *On²Vers* services in the eScience network
 - So far: temporary user management
 - Combine with eScience platform (eScience users \supseteq *On²Vers* users?)

DISCUSSION, POSTER & DEMO

Today, 13:30





Network
eScience



HTWK
Leipzig
UNIVERSITÄT LEIPZIG

On²Vers

Online Ontology Versioning

Anika Groß (gross@informatik.uni-leipzig.de)
Abteilung Datenbanken, Institut für Informatik, Universität Leipzig
Hochschule für Technik, Wirtschaft und Kultur (HTWK) Leipzig

Europa fördert Sachsen.



ESF
Europäischer Sozialfonds



SACHSEN

Motivation

- Ontologies**
- Structured representation of knowledge
 - Various applications: ontology-based semantic search, queries, annotations, ...
 - Very large ontologies, e.g. in the life sciences



- Ontology Evolution**
- Ongoing research, new findings
 - continuous modifications of ontologies
 - Periodical release of new ontology versions
 - Impact on dependent data and applications!



- Aims**
- Provide services for Online Ontology Versioning (On²Vers) in the eScience network
 - Efficient version management, multi user application
 - Basic and complex evolution analyses, mapping migration

Versioning and Change Monitoring

- User specific & "open" projects
- Integration of ontology versions into GOMMA repository
- Application of services for basic and complex evolution analyses (CODEX)
- Examples: Visualization Ontology (VISO), Adult Mouse Anatomy Ontology

Version Import

Basic Change Monitoring

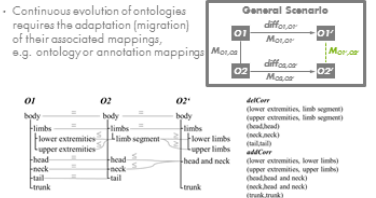
Complex Diff

System Overview



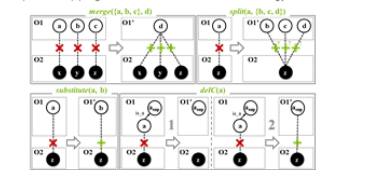
Mapping Migration

- Continuous evolution of ontologies requires the adaptation (migration) of their associated mappings, e.g. ontology or annotation mappings



Semi-automatic mapping migration

- Update mappings based on the diff between ontology versions



- Future work**
- Integrate Mapping Migration Service and Region Explorer into On²Vers
 - Provide On²Vers services in the eScience network

References

- Groß, A.; Das Reth, J.C.; Harung, M.; Fuks, C.; Rahm, E.: SemiAutomatic Adaptation of Mappings between Life Science Ontologies. Proc. 9th Int. Conference on Data Integration in the Life Sciences (DILS), 2013.
- Harung, M.; Groß, A.; Rahm, E.: CODEX: Evaluation of semantic changes between ontology versions. Bioinformatics 28 (6), 2012.
- Kruse, T.; Groß, A.; Harung, M.; Rahm, E.: GOMMA: A Component-based Infrastructure for managing and analyzing Life Science Ontologies and their Evolution. Journal of Biomedical Semantics 2011.
- Visualisation Ontology (VISO): <http://purl.org/viso/>